14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

Hard Sensor Fusion for COIN Inspired Situation Awareness

Matthew S. Baran
Donald J. Natale
Richard L. Tutwiler
Communication and Imaging Division
Applied Research Laboratory
The Pennsylvania State University
University Park, PA, USA
{msb236,natale,rlt1}@psu.edu

Christopher Griffin
John M. Daughtry
Cyber Innovation Division
Applied Research Laboratory
The Pennsylvania State University
University Park, PA, USA
{cxg286,daughtry}@psu.edu

Matthew M. McQuillan
NSWC-Dahlgren Division
6096 Tisdale Road-Suite 308
EO-IR Sensor Technology & Development G83
Dahlgren, VA 22448-5156
matthew.mcquillan@navy.mil

Jeffrey Rimland
David L. Hall
College of Information Sciences & Technology
The Pennsylvania State University
University Park, PA, USA
{jrimland,dhall}@ist.psu.edu

Abstract - Counter Insurgency operations require the ability to develop accurate representations of the physical environment and the human landscape in various conditions (e.g., urban and non-urban, day and night, and various weather conditions). We are developing innovative sensor suites and processing techniques suitable for such domains as part of a larger effort to support human-centric hard/soft data fusion. In this paper, we present a sensor suite, an information processing architecture, examples of the resulting fused information, and future experimental designs. These combined resources present opportunities for creating rich 3-D characterizations of the environment and can support novel hybrid human/computer methods for target characterization, identification, and tracking.

Keywords: counter insurgency, hard data, sensor fusion, LIDAR, SWIR, MWIR, 3D, synthetic data.

1 Introduction

Counter Insurgency (COIN) operations require the ability to develop accurate representations of the physical environment (e.g., terrain, vegetation, buildings, roads, vehicles, etc.) as well as the human landscape (individuals, crowds, etc.). Such operations potentially involve both urban and non-urban environments and must be conducted in various day/night and weather conditions.

Mixed hard and soft sensor fusion is an arising challenge for the fusion community. Over the past 3 decades, we have made substantial progress in hard sensor fusion focusing primarily on one-dimensional signals and the analysis of kinematic data [1-5]. More recent work has also focused on the fusion of logical or non-signals based data with some success [6-8]. At the same time, there has been an explosion of results on natural language processing ([9-13] are but a small sample of this work). Integration of these two disparate fields to create a common framework for the analysis of COIN operations may represent a fundamental shift in our prosecution of the global war on terror and peace keeping operations in slowly stabilizing countries. In particular it could potentially allow us to integrate and understand the myriad of sensor information collected from in situ and stand-off sensor systems with the perpetually updated information collected from the Internet via social network sites or other covert collection operations.

Several research questions arise as a result of the hard and soft sensor fusion problem. Automated extraction of entities and geo-intelligence is required to task responders for action as a result of text and image information. Co-registration of geographic features in images with those described in text is a new challenge for the sensor fusion community. Historical veracity of both soft and hard sensor feeds must be addressed before tasking results from fusion products. Co-verification from orthogonal sources like hard and soft sensors represents a unique avenue for advancement in deception detection. Processing text for highly ambiguous terms may require human-interaction in the absence of substantial training data sets, which may not be available. These problems may be solvable by including concepts from human computation [e.g., 14] in which a fusion system reduces the number of hypotheses available and a human completes the fusion process. Detection of events (verbs) inside audio / video streams is complex. Integration of audio / video processing with text / sensing making will

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2011		**00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE	
Hard Sensor Fusion for COIN Inspired Situation Awareness	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Pennsylvania State University,Communication and Imaging Division,Applied Research Laboratory,University Park,PA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Presented at the 14th International Conference on Information Fusion held in Chicago, IL on 5-8 July 2011. Sponsored in part by Office of Naval Research and U.S. Army Research Laboratory.

14. ABSTRACT

Counter Insurgency operations require the ability to develop accurate representations of the physical environment and the human landscape in various conditions (e.g., urban and non-urban, day and night, and various weather conditions). We are developing innovative sensor suites and processing techniques suitable for such domains as part of a larger effort to support human-centric hard/soft data fusion. In this paper, we present a sensor suite, an information processing architecture, examples of the resulting fused information, and future experimental designs. These combined resources present opportunities for creating rich 3-D characterizations of the environment and can support novel hybrid human/computer methods for target characterization, identification, and tracking.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

be key. Verb detection in images is challenging, especially in compressed video.

Currently, we are focused on the question of devising an over-arching architecture for the integration of hard and soft sensor fusion and the specific sub-architectural features necessary for hard sensor fusion of multimodal image information. In the following sections, we present the design of our sensor suite, the information processing architecture, COIN operations verification and validation scenarios, and the derived hard fusion. We conclude by outlining future work with respect to both the current work and the larger project.

2 Sensor Suite

Four criteria drove our hard sensor suite design (Figure 1). First, we wanted our sensor suite to have ecological validity. By this, we mean that we wanted to select sensors representative of tactically deployed sensors [e.g., 17]. Second, we only selected sensors which provide informational "value added" to the inference process for our selected targets. Third, we sought out sensors which could be demonstrated; namely, those which could be utilized in real demonstrations and campus-based experiments. Fourth, we wanted at least one sensor that could allow for innovation in the hard sensor fusion processing flow.

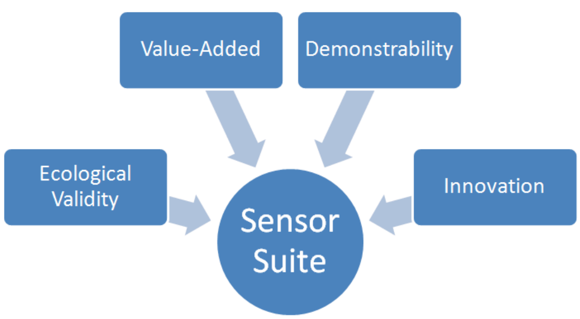

Figure 1: Sensor Suite Design Criteria

The result of this selection was the following suite: Light Detection and Ranging (LIDAR) which operates in the Short-Wavelength Infrared (SWIR) band, combined with Mid-Wavelength Infrared (MWIR), visual video, and acoustic sensors. The fusion applications that will be demonstrated in this paper combined the Flash LIDAR with a MWIR sensor. 3D Flash LIDAR uses an array of independent LIDAR receivers and focuses light returning from the scene onto the array with a lens system. In many ways it has a user aesthetic which is exactly like any ordinary digital video camera. The flash is generated by an on-board laser module and a beam-spreading optical element. This is analogous to flash photography with conventional 2D digital video cameras. Since all pixels function in parallel to each other the motion of the platform and motion of the scene between pixels samplings is zero. The 3D flash LIDAR camera chosen generates relatively noise-free point cloud videos at ranges up to 1.5 kilometers and at frame rates up to 30 Hz. The laser wavelength of the camera is 1.57 μm and is considered eye-safe. The laser is pulsed for only 5 ns per video frame. The entire camera unit is approximately 11 × 6 × 6 inches in size, weighs approximately 10 lbs and can be powered by a standard 110V wall outlet or by a 12V motorcycle battery. Mid-wavelength infrared (MWIR, IR-C *DIN*) is also called intermediate infrared (IIR): 3-8 µm. The 3 to 5 micron band is defined by the atmospheric window and covered by Indium antimonide [InSb] and HgCdTe and partially by lead selenide [PbSe]). In guided missile technology the 3-5 µm portion of this band is the atmospheric window in which the homing heads of passive IR 'heat seeking' missiles are designed to work, homing on to the IR signature of the target aircraft, typically the jet engine exhaust plume. The assembled system is illustrated in Figure 2 and the specifications are listed in Table 1.

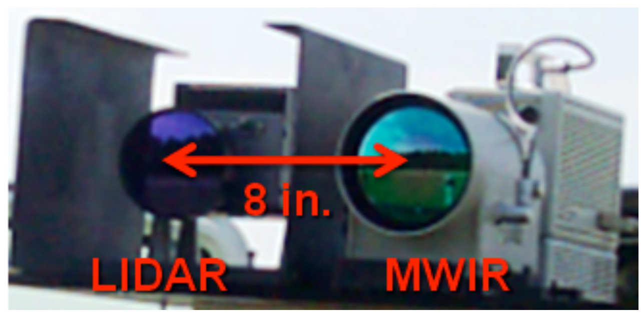

Figure 2: Sensor Suite

Table 1: Sensor Specifications

Sensor	FOV (deg)	Pixel Pitch (um)	FPA Dimensions	Frame Rate (Hz)
Flash Lidar (SWIR)	3	100	128 x 128	20
MWIR	2.23	30	256 x 256	60

3 Information Processing Architecture

Our proposed information flow architecture is shown in Figure 3. Both hard and soft sensor information is collected and can be individually fused to create an intermediate fusion product. In the case of soft sensing, this may be combined by key phrase detection, entity extraction or common text area identification. In the case of hard sensing this may be fused LIDAR and infrared images or time synchronization of video with still camera information.

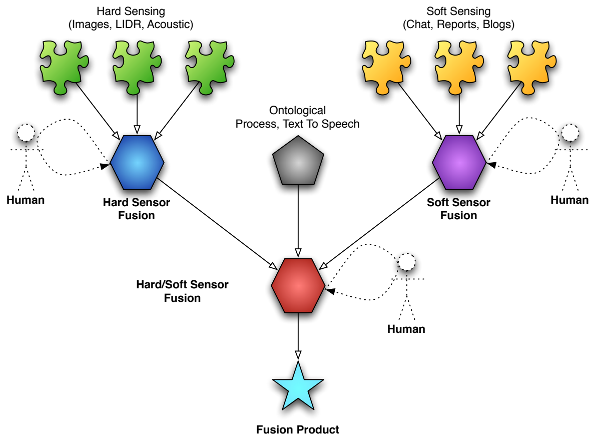

Figure 3: Proposed hard / soft sensor fusion architecture

These two fusion products can then be fed into a hard / soft sensor fusion system. The integration of soft with hard fused data will require the use of an ontological processing entity to allow verbs or nouns in the text to be identified within the images [15, 16]. Unfortunately, the inclusion of ontologies may make the system substantially more brittle and difficult to update. Therefore, a looser extensible dictionary of terms that can be detected in images may be substantially more effective than a full ontology. Humans-in-the-loop may be present at all levels of the fusion activity. These humans can act as proxies for algorithms or can serve to disambiguate algorithm confusion associated with e.g., ambiguous terms in text analysis or corrupt images in hard sensor fusion. In the following sections, we focus on the fusion of mixed image intelligence data, thereby providing details behind the components that might function in the left-hand-side (hard sensor fusion component) of the proposed fusion architecture.

4 Sensor Fusion Scenarios

To support the development and testing of our sensor suite, information processing architecture, and data fusion capabilities, we have developed a series of scenarios that realistically simulate both urban environments (in which targets might be obscured by walls, smoke, crowds, etc) and wooded environments (where thick foliage can interfere with sensing and observations).

4.1 Scenario 1 – Urban

In the first scenario, a simulated urban area is populated with moving vehicles and simulated innocent bystanders and pedestrians. The "Blue team" of agents patrols the area and reports their findings using communication devices (COMM). The Blue team also uses a Flash LIDAR sensor bore-sighted with a MWIR sensor with the capability to pan and tilt as needed.

Figure 4: Urban Scenario Environment

The "Red team" consists of a group of individuals who drive to a building (Building 1 in Figure 5) with a central location within a populated area. Two of the individuals on the Red team leave the vehicle and proceed to take positions in second-story windows in that building. The Red team vehicle then proceeds to a nearby location (behind Building 2 in Figure 5) that people would be likely to flee to in order to take cover in the event of a shooting. While this occurs, the data feeds from the LIDAR, MWIR, and COMM channels are being monitored and recorded by the (Blue team) data fusion system. The two Red team individuals in window locations in the building open fire on the crowd below. When this occurs, the crowds take cover near the abandoned Red team vehicle, which is then detonated.

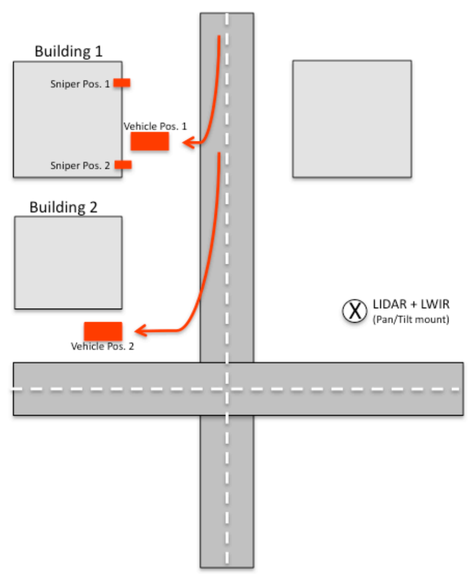

Figure 5: Urban Scenario

The scenario is described above in the nominal case as it is intended to occur by the Red team. In certain instances, it can play out differently if the data fusion system is able to detect the threat and recommend an alternate course of action for the Blue team or direct a human analyst's attention to information that would help them arrive at a preferable Blue team action plan.

This scenario includes the potential for challenges such as dense smoke to impede vision, crowds of innocent people to make identification of Red team members more difficult, and vehicular traffic to add additional complexity. This, combined with realistic simulated radio COMM, provides an ideal opportunity to exploit the capabilities of LIDAR technology and hard/soft sensor fusion techniques.

4.2 Scenario 2 – Dense Vegetation

In the second scenario, we simulate a planned improvised explosive device (IED) ambush attack in a densely wooded area. In this scenario, the Blue team is patrolling the area with the goal of protecting a convoy. Meanwhile, the Red team has planted a roadside IED and is planning to ambush the Blue team convoy. While this situation unfolds, fused Flash LIDAR and LWIR data feeds and COMM dialog between Blue team members are streamed to the data fusion system for analysis.

Figure 6: Dense Vegetation Environment

Much like in the urban scenario, the challenges faced by the Blue team in this scenario leverage the capabilities of LIDAR as well as the utility of hard/soft data fusion techniques. In this case, the dense vegetation of the surrounding area combined with the linear approach of the convoy along the road and the unconstrained movement of the Red Ambush Team and Blue Patrol Team provides a particularly interesting sensing and data fusion challenge.

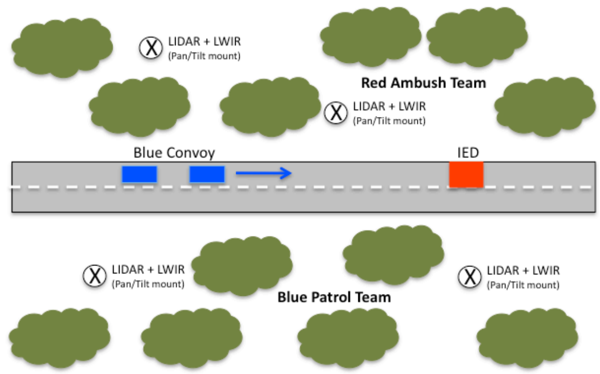

Figure 7: Dense Vegetation Scenario

5 Hard Fusion of LIDAR & IMINT

Image fusion across modalities is a challenging problem from accurate registration to meaningful representation of the fused information. A fused product must convey the important information from each modality in a way that can be naturally interpreted by a human observer. We propose a method of fusing 3D range information from a Flash LIDAR with a thermal MWIR image to convey the location of objects of interest within the focal plane and naturally within 3-space. The fusion method makes use of human visual perception of color and brightness to convey range and temperature, respectively.

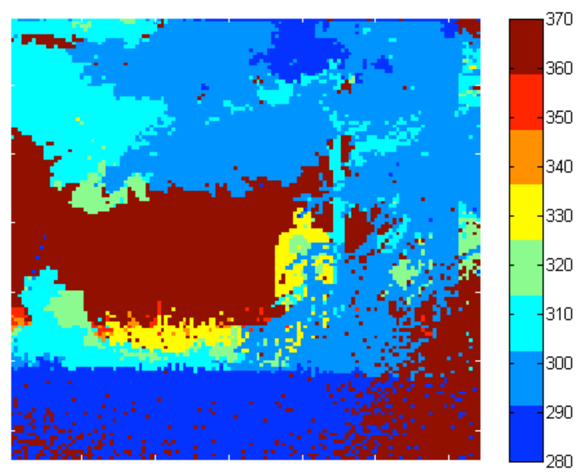

Figure 8: A Flash LIDAR range image mapped to 8 color bins within the ranges of 280 ft. to 370 ft.

Given a range image R directly mapped to ortho-rectified (x,y,z) points in 3-space, and a thermal image T, assumed to be registered pixel-by-pixel to the range image, a fused image may be constructed using data from each source. Given a colormap M_R with n bins, $M_R\colon R \rightarrow R_n$ divides the range image into n colors where $R_n \in \mathbb{R}^{px3}$ for p pixels (Figure 8). Each row of R_n provides a red, green, and blue value to define the pixel color. Given an intensity map M_I with k bins, $M_I\colon I \rightarrow I_k$ discretizes the intensity image into k bins where $I_k \in \mathbb{R}^{px1}$ and each $I_k(i) \in [0,1]$, i=1,…,p (Figure 9). The Fused image F is defined as $[I_k I_k I_k] \circ M_R$, where $\circ$ denotes the entry-wise product. This scales the intensity of the pixel colors by the intensity of the thermal image (Figure 10).

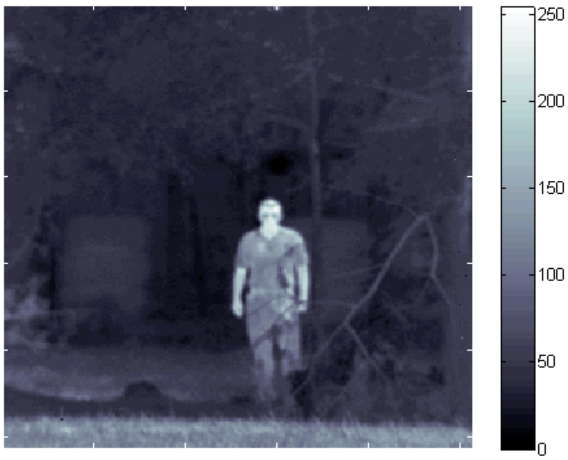

Figure 9: A MWIR thermal image mapped to 256 bins with values in [0 1]

6 Conclusions and Future Work

In this paper, we presented a hard/soft information fusion architecture, a hard sensor suite (accompanied with the rationale driving the design), COIN-inspired scenarios for driving our empirical analyses, and the hard sensor fusion algorithms and resulting human-centered information products.

As we continue with this work, we will: (1) conduct experiments centered on the scenarios described, (2) explore the algorithmic design space that arises from the resulting data, and (3) investigate approaches to the fusion of our hard and soft data sets.

The two primary locations for our planned experiments are the Penn State campus at University Park, PA and a fire safety facility located nearby. The Penn State campus is ideal for using non-threatening analogous scenarios to investigate human-in-the-loop issues such as knowledge elicitation from participating observers, dynamics of centralized vs. distributed command, motivation of human observers, and team cognition. Additionally, the Extreme Events Lab [18] located on the Penn State campus can serve as a command center for these experiments. The fire safety facility serves as an ideal location for the hard sensor experiments described above.

Data fusion is in many ways a design science. Rather than just developing algorithms, we must seek out a deeper understanding of the nature of our algorithms. Indeed, the selection of algorithmic approaches cannot be centered solely on the efficacy of a given algorithmic

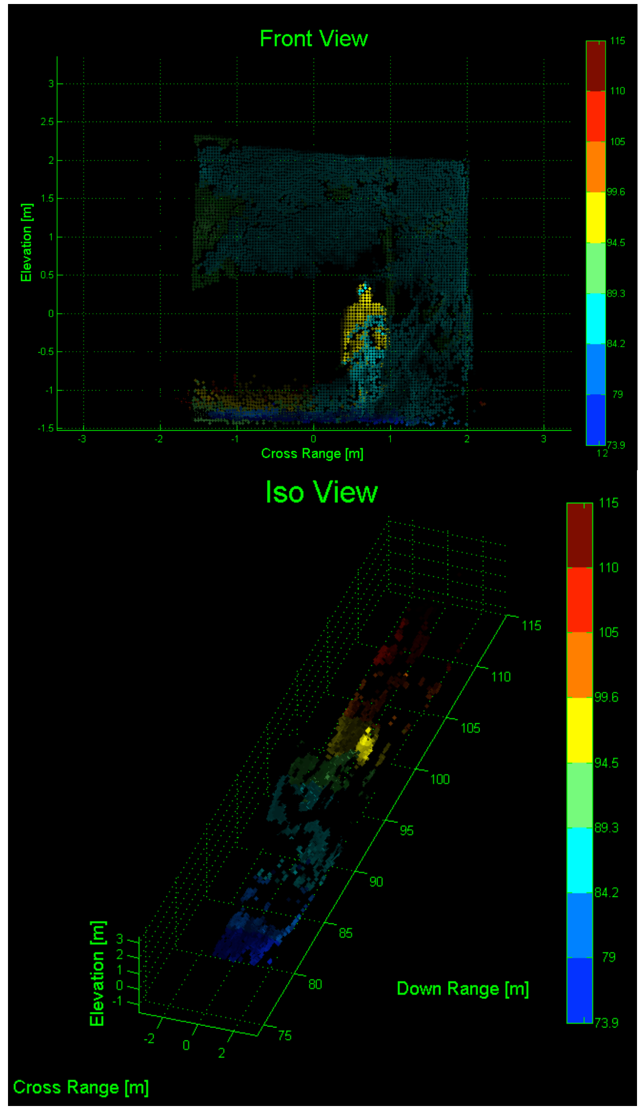

Figure 10: Two views of the fused data showing only points between 75 m. and 115 m. The human has a higher temperature than the background resulting in more vibrant colors that clearly indicate his range at 100m.

approach, but rather on the competing alternative trade-offs. How does a given algorithmic approach support, fail to support, or even undermine our goals by filling, failing to fill, or undermining the information needs – information delivery – sensemaking - action cycle?

The SYNCOIN dataset [19] has been constructed by Penn State researchers in order to provide a substantial set of synthetic soft data with corresponding hard sensor data opportunities. SYNCOIN contains approximately 600 messages that represent COIN-inspired scenarios. Additionally, “ground truth” and pedigree metadata documents are maintained for all messages and their corresponding threads of interest. The activities and experiments described in this paper will complement SYNCOIN by providing relevant hard sensor datasets.

This combination of a new sensor suite, realistic synthetic hard and soft data, relevant metadata/ground truth documents, and the capability to perform human-in-the-loop experimentation represent an opportunity to break new ground in human-centric hard and soft information fusion.

7 Acknowledgements

We gratefully acknowledge that this research activity has been supported in part by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for "Unified Research on Network-based Hard/Soft Information Fusion", issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery.

References

[1] R. R. Brooks and S. S. Iyengar, *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice Hall, 1997.

[2] M. E. Liggins, D. L. Hall, and J. Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC Press, 2 ed., 2008.

[3] D. Hall and S. A. H. McMullen. *Mathematical Techniques in Multisensor Data Fusion.* Artech House, 2 ed., 2004.

[4] R. R. Brooks, C. Griffin, and D. Friedlander, "Self-organized distributed sensor network target tracking," *J. High Performance Computing*, Special Issue on Sensing, vol. 16, no. 3, pp. 207–220, 2002.

[5] D. L. Hall and J. Llinas, eds., *Handbook of Multisensor Data Fusion.* CRC Press, 2001.

[6] H. Chen and P. Meer, "Robust fusion of uncertain information," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 35, pp. 578–586, June 2005.

[7] C. Griffin and S. Racunas, "Logical data fusion for biological hypothesis evaluation," in *Proceedings of FUSION 2005*, (Philadelphia, PA, USA), July 24-28 2005.

[8] C. Griffin, K. Testa, and S. Racunas, "An algorithm for searching an alternative hypothesis space," *IEEE Trans. Sys. Man and Cyber. B* (Forthcoming), November 2010.

[9] C. D. Manning and H. Schu ̈tze, *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

[10] D. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[11] T. Hofmann, "Probabilistic latent semantic indexing," in *Proceedings of the Twenty-Second Annual International SIGIR Conference on Research and Development in Information Retrieval* (SIGIR-99), 1999.

[12] B. W. Bader, M. W. Berry, and M. Browne, "Discussion Tracking in Enron Email using PARAFAC," in *Survey of Text Mining II*, Springer London, 2008.

[13] M. W. Berry, S. T. Dumais, and G. W. O'Brien, "Using linear algebra for intelligent information re-trieval," *SIAM Review*, vol. 37, no. 4, pp. 573–595, 1995.

[14] G. Little, L. B. Chilton, M. Goldman, and R. C. Miller, "Exploring iterative and parallel human computation processes." *Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP),* 2010, 68-76.

[15] F. Baader, *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, January 2003.

[16] M. Smith, C. Wetty, and D. L. McGuiness, "OWL Web Ontology Language Guide," technical report, W3C, Feb. 10 2004.

[17] Ground-Based Operational Surveillance System (G-BOSS). Retrieved 02/25/2011 from http://www.marcorsyscom.usmc.mil/sites/cins/Fact%20Books/INTEL_MISC/G-BOSS_OIF.pdf]).

[18] D. Hall, J. Graham, L. More and J. Rimland, "Test and evaluation of soft/hard information fusion systems: an experimental environment, methodology and initial data sets" in *Proceedings of the 13th International Conference on Information Fusion,* Edinburgh, UK, July, 2010

[19] Graham, Rimland, & Hall, "A COIN-inspired synthetic dataset for qualitative evaluation of hard and soft fusion systems." in *Proceedings of the 14th international conference on information fusion*. Chicago, IL. 2011 (Manuscript accepted)